

LA ESCALA KEYMATH-R: ESTUDIO DE UNA PRUEBA DIAGNÓSTICA CON REFERENCIA CURRICULAR

Ricard Marí Mollà
Universitat de València¹

RESUMEN

Las nuevas tendencias de la investigación exigen a los instrumentos de evaluación y diagnóstico una adecuación a las nuevas orientaciones de la enseñanza y una fundamentación en los propios procesos de aprendizaje. En el caso del aprendizaje de las matemáticas y desde un enfoque criterial, se presenta una revisión de la escala KeyMath-R, una prueba referida a los contenidos o dominios, que ha mostrado una organización interna correspondiente al currículum, con unos niveles de bondad métrica muy aceptables, considerándose un instrumento diagnóstico de gran utilidad.

ABSTRACT

The new investigation trends require to the evaluation and diagnostic tools a learning new trainings adaptation and foundation in the own learning process. For the mathematics achievement evaluation and with criterion-referenced, is made a KeyMath-R scale review. This is a domain-referenced inventory, with an internal organization corresponding to the curriculum. The reliability and validity coefficient levels show enough acceptability, which support its use.

Desde hace más de una década el nombre de **KeyMath-R** es un sinónimo de investigación educativa y está presente en estudios sobre tests de rendimiento y en los listados de pruebas educativas. Es una escala para evaluar las habilidades y capacidades

¹ Ricard Marí Mollà es Profesor Titular de la Universitat de València, adscrito al Departamento MIDE. E-meil: Ricard.Mari@uv.es

matemáticas y particularmente útil para el diagnóstico de los déficits académicos en sujetos con dificultades de aprendizaje y, además, por su adecuación curricular permite diseñar programas de intervención educativa.

Su éxito se deriva de su alcance comprensivo, de su contenido motivador y colorista, y de la facilidad de su administración, puntuación y aplicación a la práctica instruccional. El elemento principal lo integra el *KeyMath Revised: A Diagnostic Inventory of Essential Mathematics* (Connolly, 1988).

El KeyMath-R forma parte de un conjunto de varios elementos: un cuestionario diagnóstico en sus formas A y B, un suplemento métrico a la prueba, un programa de desarrollo del primer nivel, un programa de rehabilitación y un manual de actividades manipulativas y ejercicios (Connolly, 1985, 1988 y 1988a).

El **KeyMath-R** es una segunda construcción de una prueba de evaluación de logros de aprendizaje en matemáticas elementales realizada por A.J. Connolly, el *Key-Math*, en este caso mejorada notablemente en sus aspectos de amplitud de aplicación, de adaptación curricular y normalización, para que pueda ser utilizado con todo tipo de sujetos, normales o con dificultades de aprendizaje y en diferentes situaciones de aprendizaje, tanto para la evaluación en el aula normal como en el ámbito del diagnóstico de las dificultades de aprendizaje y para todos los niveles de la escuela primaria.

Está considerado como un «Inventario Diagnóstico de Matemáticas Básicas». Una escala de aplicación individual que evalúa las habilidades matemáticas correspondientes a los diferentes niveles de la educación primaria y que consta de 13 Subpruebas clasificadas en 3 Subescalas: Conceptos Básicos, Operaciones (procesos de cálculo) y Aplicaciones. Es una prueba con referencia a los contenidos o dominios y validado para la interpretación normativa. Según Bachor (1989-90) en diferentes estudios sobre el KM-R se le ha considerado por los autores como test normativo, por las tablas normativas del manual, como test referido al dominio e, incluso, como test referido al criterio (Tinney, 1975; Price, 1984).

El interés de la prueba, aparte de su validez demostrada durante muchos años en los E.E.U.U. en el diagnóstico de las dificultades de aprendizaje de las matemáticas, tanto en su construcción original como en la revisión que estamos analizando, radica en que identifica en cada momento el contenido esencial (conceptos y destrezas) del trabajo del alumno, organizado en un marco conceptual que facilita la realización del diagnóstico descriptivo y la planificación de la futura instrucción.

Se trata de un instrumento de administración individual, que entre 35 y 50 minutos puede ofrecer una evaluación comprensiva del conocimiento y aplicación de conceptos o habilidades matemáticas importantes de un alumno. Si bien la intención del KM original estaba clara, como se ha dicho, el propósito del KM-R no ha sido definido por el autor, no se sabe si su diseño está pensado para los alumnos especiales o los normales. No obstante, siguiendo a Bachor (1989-90), según la selección de sujetos efectuada en el proceso de estratificación de las variables siguiendo el diseño de validación, se considera adecuado para ser utilizado tanto con sujetos considerados normales, como con sujetos afectados por variables raciales o étnicas que supongan deficiencias culturales y/o lingüísticas.

ELEMENTOS DE LA ESCALA KEYMATH-R

Alcanza un rango amplio de edades: desde Preescolar, hasta el noveno grado. Dispone de dos formas alternas, A y B, equiparadas estadísticamente y por contenido. Su contenido se basa en un sistema comprensivo referido al dominio que identifica jerarquías de conceptos y habilidades a través de trece subpruebas de contenido. Cada una de las 13 subpruebas del sistema se mide por separado. Las subpruebas, por su parte, se agrupan en áreas de contenido matemático y evalúan dicho contenido en las tres subescalas ya mencionadas de: Conceptos Básicos, Operaciones y Aplicaciones (Ver Tabla 1).

Tabla 1
ESTRUCTURA DE CONSTRUCCIÓN DE LA ESCALA KM-R

Conceptos Básicos	Operaciones	Aplicaciones
Numeración	Suma	Medida
Número Racional	Resta	Tiempo y Dinero
Geometría	Multiplicación	Estimación
	División	Interpretación datos
	Cálculo mental	Resolución problemas

Cada subprueba contiene tres o cuatro dominios con una importancia instruccional similar. Cada uno de estos dominios se representa por un conjunto de seis ítems. Con esta representación el KM-R puede ofrecer datos fiables por debajo del nivel de subprueba, los cuales, según el constructor, facilitan el análisis y la traslación de los resultados de la prueba a alternativas instruccionales.

Ofrece información diagnóstica a cuatro niveles: Total Escala, Subescalas, Subpruebas y Dominios. Cada nivel ofrece una información específica sobre la ejecución del alumno al identificar los puntos fuertes y débiles, es decir, lo que sabe y lo que no sabe el alumno y, de esta forma, se establece una base para confeccionar un programa instruccional adecuado:

Hay que añadir, que, según las intenciones del autor, las interpretaciones de los resultados de la prueba hechas en los niveles ya mencionados, se pueden aumentar con un análisis de la ejecución del estudiante sobre ítems específicos, los cuales aparecen ordenados de menor a mayor grado de dificultad en cada uno de los dominios. Para esta finalidad, el autor proporciona, además de los ítems, un listado de los objetivos de los mismos.

CONTENIDO DE LA ESCALA KM-R

Teniendo en cuenta la exposición anterior y la definición que del mismo realiza A.J. Connolly en el manual, el KeyMath-R es una prueba referida al contenido y con provisiones para una interpretación normativa. A través de su desarrollo, el énfasis se sitúa en la identificación del contenido esencial (conceptos y habilidades) y en la organización de dicho contenido en un sistema que facilite la evaluación diagnóstica y la planificación de la futura instrucción.

En su construcción, Connolly tuvo que resolver un importante problema derivado de la búsqueda del contenido matemático esencial. A partir de la revisión de los programas matemáticos básicos se encontró con que no existía un modelo consensuado para la organización y delineación del contenido matemático en la enseñanza elemental y media. Esta dificultad se complicaba todavía más por los problemas que se presentaban al intentar unir los esfuerzos en la evaluación y rehabilitación con el desarrollo instruccional.

Así pues, la construcción del sistema de contenido la empezó Connolly con una amplia revisión de los currícula matemáticos existentes, de las prioridades instruccionales actuales y de las tendencias futuras. Este esfuerzo incluyó una revisión del alcance y secuencia de algunos libros de texto básicos de la enseñanza elemental y las recomendaciones curriculares establecidas en *An Agenda for Action* del National Council of Teachers of Mathematics (1980).

Posteriormente este problema se ha resuelto de alguna forma al imponerse las recomendaciones del NCTM en el área de las matemáticas elementales y configurarse como un estándar para los educadores (NCTM, 1988 y 1989; Tej Pandey, 1994).

La revisión de la escala KeyMath-R realizada por nosotros (Marí, 1996) incluye sendos trabajos de Validación de Contenido desde la perspectiva del Diseño Curricular tanto del estándar del NCTM como de las prescripciones para la implantación de la LOGSE, que completan y mejoran el proceso de estudio del contenido y la Validez de Constructo de la adaptación original de la escala.

ALCANCE Y SECUENCIA

El sistema de la escala KM-R comprende, pues, tres *Subescalas*: Conceptos Básicos, Operaciones y Aplicaciones, a las que se les concede la misma importancia instruccional. A su vez, las tres *Subescalas* anteriores se divide en trece *Subpruebas*, seleccionadas y desarrolladas con una importancia equivalente, y cada una de ellas, a su vez, en tres o cuatro *Dominios*, que de nuevo mantienen el principio de equilibrio en la consideración de la importancia entre los diferentes dominios.

La especificación del contenido cuidadosamente articulada y equilibrada, descrita en la Tabla siguiente (ver Tabla 2), representa la aplicación desde los currícula en la que el valor instruccional varía marcadamente a lo largo de los subpruebas o a lo largo de los dominios.

Tabla 2

ESPECIFICACIÓN DEL CONTENIDO DE LA ESCALA KEYMATH-R: SUBESCALAS, SUBPRUEBAS Y DOMINIOS

Subescala:	CONCEPTOS BÁSICOS	OPERACIONES	APLICACIONES
Subprueba: Dominios:	<p>Numeración Números 0-9 Números 0-99 Números 0-999 Números multidígitos y tópicos de numeración avanzados</p> <p>Números Racionales Fracciones Decimales Percentiles</p> <p>Geometría Relaciones espaciales y atributivas Formas bidimensionales y sus relaciones Geometría coordinada y transformacional</p> <p>Formas tridimensionales y sus relaciones</p>	<p>Suma Modelos y hechos básicos Algoritmos para sumar números enteros Suma de números racionales</p> <p>Resta Modelos y hechos básicos Algoritmos para restar números enteros Resta de números racionales.</p> <p>Multiplicación Modelos y hechos básicos Algoritmos para multiplicar números enteros Multiplicación de números racionales comerciales</p> <p>División Modelos y hechos básicos Algoritmos para dividir números enteros División de números racionales</p> <p>Cálculo mental Cálculos encadenados Números enteros Números racionales</p>	<p>Medida Comparaciones Uso de unidades no estándar Uso de unidades estándar: longitud, y área Uso de unidades estándar: peso y capacidad</p> <p>Tiempo y dinero Identificación del paso del tiempo Uso de relojes y de las horas Uso monetario de un dólar Uso monetario de cien dólares y transacciones</p> <p>Estimación Números racionales y enteros Medidas Cálculo</p> <p>Interpretación de Datos Cuadros y tablas Gráficos Probabilidad y estadísticos</p> <p>Solución de problemas Resolver problemas rutinarios Comprensión de problemas no rutinarios Resolver problemas no rutinarios</p>

El KM-R es, pues, un instrumento de evaluación basado en un sistema específico, cada apartado de conocimiento de las tres subescalas del KM-R se mide por una subprueba separada. Cada dominio se mide por un conjunto de seis ítems, a través de los que se nos da el nivel alcanzado por el estudiante en la realización de la prueba. El rigor aplicado a la selección y determinación del contenido, expuesto anteriormente, se empleó, también, en la formulación y desarrollo de los ítems. Cada ítem fue sometido a un estudio, consulta de expertos y demostraciones apropiadas de sus características psicométricas. Los ítems que superaron estas pruebas son los que se presentan en las diferentes subpruebas del KM-R distribuidos secuencialmente, no por los dominios que representa y pretende evaluar cada uno de ellos, sino siguiendo un orden de dificultad previamente establecido.

ESTUDIO MÉTRICO DIFERENCIAL DE LA ESCALA KEYMATH-R

La necesidad de disponer de instrumentos diagnósticos que atiendan a la totalidad tanto de los contenidos como de los elementos procesuales del aprendizaje de las matemáticas, así como poder disponer de instrumentos diagnósticos que proporcionen elementos para la intervención pedagógica y superen las deficiencias de los tests normativos al uso, ha aconsejado el estudio de algunas pruebas orientadas a la evaluación diagnóstica. El KeyMath-R es uno de dichos instrumentos diagnósticos. Tal como se ha señalado, por su referencia a los contenidos o dominios, las descripciones de los campos de conducta a evaluar, la adecuación curricular de los dominios propuestos en línea con las novedades curriculares, su impacto de uso en el campo diagnóstico, los estudios sobre criterios de bondad y el volumen de investigación desarrollado sobre el mismo, lo hacen un instrumento valioso, por lo que se ha decidido un análisis en profundidad desde nuestro ámbito cultural.

En primer lugar, dado que la escala KM-R sólo muestra una referencia curricular y para dotarla de una mayor riqueza diagnóstica, se ha realizado una revisión teórica de algunas de las conclusiones aportadas por la investigación desde el enfoque del Procesamiento de la Información y referidas a los procesos de la enseñanza y aprendizaje de las matemáticas en la escuela primaria. Con ello dotamos de una fundamentación teórica a la escala *KeyMath-Revised*, la cual otorga a la prueba una mayor riqueza de interpretación diagnóstica y permiten enriquecer los programas de intervención con las conclusiones de la propia psicología cognitiva.

En segundo lugar, se realiza un estudio empírico y de adaptación del cuestionario basado en una estructura de análisis clásico de construcción de pruebas, abordando aquellos componentes que se han considerado válidos para la depuración de un instrumento de medida: descripción de la muestra utilizada y su rendimiento, fiabilidad, validez y tipificación. Se ha tomado como referencia el modelo clásico, aunque algunos autores actualmente hacen hincapié en otros elementos como el TRI (Teoría de Respuesta al Ítem). Se ha insistido más en el ajuste del instrumento en aspectos de validez, primando elementos de análisis de contenido sobre los de depuración métrica (ej.: modelo Rasch de fiabilidad), puesto que se ha considerado de

mayor interés interpretativo dadas las características de referencia criterial de la prueba.

La **muestra** de sujetos para la adaptación del cuestionario KeyMath-R está constituida por un total de 1.238 sujetos de la Comunidad Valenciana, seleccionados al azar, de los que el 50,2% son mujeres y el 49,8% varones con un rango de edad comprendido entre los 5 y los 15 años. Todos ellos son estudiantes de los cursos comprendidos entre 2º de Educación Preescolar, EGB y 1º de BUP-1º FP.

La muestra seleccionada es representativa de la población escolar comprendida entre los niveles de Preescolar, de la EGB y 1º BUP-FP, a nivel de la Comunidad Valenciana, y se realizó sobre el informe de la población escolar de la Conselleria d'Educació i Ciència de la Comunitat Valenciana correspondiente al curso 1989-90, con el fin de asegurar una proporcionalidad estricta a la población escolar mencionada.

Para la configuración de la muestra se han tenido en cuenta diversas variables de estratificación como: curso o nivel de estudios, sexo, tipo de colegio, clase social, lengua, etc., de forma que se mantenga la distribución interna de cada grupo de edad, conservando las mismas proporciones en la muestra que en la población.

El estudio empírico incluye, preceptivamente, un **análisis de ítems**. Este análisis se ha realizado desde tres perspectivas. En primer lugar, una revisión empírica de los niveles de acierto mediante el análisis del índice de dificultad (π) a lo largo de cada Subprueba para cada uno de los cursos y el total. En segundo lugar, la comprobación de si el ordenamiento de los ítems en cada Subprueba responde a una gradación de dificultad. Y, en tercer lugar, una síntesis comparativa que nos permita encontrar la adecuación entre los niveles de dificultad observados en la revisión empírica y los niveles de logro teórico, que determina el perfil teórico previsto, derivados de la revisión lógica realizada por un grupo de profesores de matemáticas seleccionados, indicando el ciclo o nivel en que cada ítem debe estar dominado según el Diseño Curricular Base.

Se comprueba que los ítems y los dominios de la Escala están graduados con un nivel de dificultad progresiva y que los niveles de rendimiento de los estudiantes coinciden con el perfil teórico de logro.

Estos análisis permiten realizar interpretaciones diagnósticas a nivel de ítem y de dominios, pudiéndose analizar los logros de los estudiantes independientemente de la interpretación de la puntuación global de la Subprueba.

Para la determinación de la consistencia de las puntuaciones obtenidas en el pase de las Subpruebas de la Escala KeyMath-R, se ha realizado un estudio de **fiabilidad**, mediante los coeficientes α de consistencia interna (Cronbach KR-20, teniendo en cuenta que son dicotómicos) y el procedimiento de las dos mitades (Spearman-Brown), para cada Subprueba y para cada Dominio, en cada uno de los niveles educativos estudiados. Los resultados obtenidos pueden considerarse buenos, con unos coeficientes mayores que, en general, a .80 y superiores a los obtenidos por Connolly en la Escala original, comprobándose la correspondencia de los coeficientes con el nivel de dificultad y con los niveles de instrucción-logro. Este hecho se considera una manifestación lógica de la adecuación de la prueba al Diseño Curricular.

Los diferentes coeficientes de fiabilidad referidos a las Subescalas, Subpruebas y Dominios, teniendo en cuenta los efectos que la variabilidad produce sobre los mis-

mos, presentan unas secuencias evolutivas que interpretamos relacionadas con la instrucción de los estudiantes.

Como ejemplos presentamos los coeficientes de fiabilidad referidos a las Subpruebas y Áreas (ver Tablas 3 y 4) en los que se observa la secuencia mostrada por los coeficientes destacando los cursos inferior (Preescolar) y superior (1º BUP), con unos niveles sin instrucción o con instrucción completa, que muestran, por ello, los coeficientes más bajos. Todo ello se considera la manifestación lógica de la adecuación de la prueba al Diseño Curricular.

Los análisis sobre la validez necesarios para la adaptación de la escala KeyMath-R, han versado sobre sendos estudios de **Validez de Constructo** y **Validez de Criterio**.

Dadas las características del KeyMath-R como prueba de rendimiento y referida al currículum, los análisis de la Validez de Constructo se han interpretado desde la perspectiva del Diseño Curricular y en la búsqueda de Evidencias de Validez de Contenido y de Validez Curricular necesarias para una interpretación de tipo criterial, mediante unos procedimientos más potentes que los utilizados por Connolly en la edición original. Así, se ha realizado una revisión de trabajos sobre validación de contenido, cuyo criterio de validación se ha centrado en la correspondencia de los dominios de la prueba con el currículum. En primer lugar, y siguiendo la estrategia de Romberg

Tabla 3

*COEFICIENTES DE FIABILIDAD S-B DE LAS SUBPRUEBAS DEL KM-R
CORRESPONDIENTES A LOS DIFERENTES CURSOS DE NUESTRA MUESTRA*

	Coeficientes de Fiabilidad Spearman-Brown										
	Preesc.	1º	2º	3º	4º	5º	6º	7º	8º	1ºBUP	1ºFP
Subprueba 1	.83	.91	.91	.89	.89	.92	.86	.80	.86	.32	.94
Subprueba 2	.74	.75	.90	.84	.93	.89	.91	.90	.92	.82	.92
Subprueba 3	.81	.90	.91	.89	.87	.87	.86	.81	.89	.73	.91
Subprueba 4	.87	.92	.87	.85	.96	.74	.76	.77	.73	.43	.68
Subprueba 5	.53	.88	.89	.85	.94	.82	.76	.82	.78	.51	.72
Subprueba 6	.56	.73	.91	.80	.73	.79	.76	.80	.77	.66	.73
Subprueba 7	.59	.69	.96	.89	.81	.92	.79	.89	.85	.82	.93
Subprueba 8	—	.83	.88	.87	.92	.94	.88	.90	.88	.75	.92
Subprueba 9	.86	.91	.84	.85	.95	.90	.82	.84	.86	.64	.85
Subprueba 10	.16	.91	.92	.91	.97	.91	.90	.84	.86	.64	.85
Subprueba 11	.52	.69	.77	.72	.92	.83	.89	.84	.79	.80	.70
Subprueba 12	.76	.84	.85	.78	.93	.90	.87	.83	.86	.67	.70
Subprueba 13	.74	.82	.79	.79	.86	.88	.88	.86	.86	.67	.78

Tabla 4

COEFICIENTES DE FIABILIDAD α DE LAS SUBPRUEBAS DEL KM-R
CORRESPONDIENTES A LOS DIFERENTES CURSOS DE NUESTRA MUESTRA

	Coeficientes de Fiabilidad α											
	Preesc.	1°	2°	3°	CURSOS							1°BUP
				4°	5°	6°	7°	8°				
Subprueba 1	.70	.86	.88	.86	.87	.86	.81	.80	.86	.28	.61	
Subprueba 2	.62	.66	.80	.73	.82	.87	.87	.88	.89	.78	.85	
Subprueba 3	.76	.83	.87	.84	.80	.85	.82	.77	.83	.73	.90	
Subprueba 4	.78	.86	.83	.71	.92	.67	.66	.67	.64	.41	.61	
Subprueba 5	.61	.82	.85	.81	.91	.74	.70	.81	.74	.58	.68	
Subprueba 6	.56	.73	.91	.80	.73	.79	.76	.80	.77	.66	.73	
Subprueba 7	.59	.69	.96	.89	.81	.92	.79	.89	.85	.82	.93	
Subprueba 8	.52	.71	.83	.82	.85	.88	.84	.84	.83	.69	.89	
Subprueba 9	.79	.89	.83	.82	.93	.87	.77	.81	.82	.65	.77	
Subprueba 10	.32	.84	.89	.88	.96	.89	.87	.80	.87	.63	.80	
Subprueba 11	.49	.65	.72	.70	.81	.84	.86	.86	.81	.75	.66	
Subprueba 12	.66	.78	.83	.80	.89	.85	.84	.85	.84	.56	.70	
Subprueba 13	.58	.76	.75	.78	.84	.86	.83	.82	.82	.70	.78	
Área 1	.53	.69	.78	.66	.81	.80	.84	.77	.86	.76	.91	
Área 2	.68	.81	.88	.83	.78	.87	.83	.86	.86	.83	.72	
Área 3	.72	.86	.83	.85	.88	.90	.91	.88	.91	.86	.85	

(1992), se ha realizado un análisis de la correspondencia de los dominios del KM-R con el currículum del NCTM de 1989. Este trabajo ha mostrado que el KM-R es más adecuado que el resto de tests de matemáticas analizados para evaluar los contenidos, procesos y niveles de pensamiento determinados por los estándares de matemáticas del NCTM y, además, ha proporcionado un criterio de validez relativo a su adecuación de contenidos en su contexto de origen.

En segundo lugar, y con referencia al contexto educativo en el que la escala KM-R ha sido adaptada, se ha presentado un segundo estudio de Validación de Contenido, basado en el análisis de la correspondencia de los ítems al currículum, representado por unos modelos de Diseño Curricular en el ámbito de la aplicación de la LOGSE y por los Programas Renovados de la EGB. Los expertos seleccionados contrastaron los objetivos conductuales de cada uno de los ítems del KM-R con los objetivos de aprendizaje implícitos en cada uno de los dominios de los Diseños curriculares propuestos indicando, para cada una de las Subpruebas, el nivel escolar en el que se inicia la instrucción del objetivo y el nivel de logro de su dominio.

Los resultados de los análisis efectuados ponen de manifiesto la existencia de dicha adecuación curricular, confirmando la intencionalidad constructiva de la prueba y aportando una evidencia importante de validez de contenido y curricular. Además, la revisión lógica de la Validez Curricular ha servido de referente teórico para los análisis efectuados y para las interpretaciones del rendimiento, confirmándose, de nuevo, el KM-R como un instrumento para la evaluación y el diagnóstico.

En tercer lugar, se ha realizado un análisis de la Estructura Dimensional de la Escala con el fin de comprobar si las dimensiones obtenidas se corresponden con los núcleos de logro de los Diseños Curriculares correspondientes a nuestro análisis teórico. Mediante el Análisis Factorial de las Subpruebas para cada curso se han elaborado una serie de conclusiones referentes a la Validez de Constructo a partir de la adecuación de las definiciones de los dominios y representatividad de los ítems del KM-R. Estas conclusiones, respecto de la estructura factorial de los totales de las Subpruebas, destacan dos factores diferenciados marcados por la existencia o no de instrucción formal (ver Cuadro 1), y, en referencia a los dominios de

Cuadro 1

ESTRUCTURA FACTORIAL DOMINANTE Y SU DEFINICIÓN, RESULTANTE DEL ANÁLISIS FACTORIAL DE LOS TOTALES DE LAS SUBPRUEBAS DEL KM-R

ESTRUCTURA FACTORIAL DOMINANTE	
Factor 1.	Subpruebas con instrucción.
Factor 2.	Subpruebas de cálculo.
<i>Factor 3.</i>	<i>Subpruebas sin instrucción.</i>

las Subpruebas, las agrupaciones responden al nivel de logro de los estudiantes antes que a la congruencia de contenido como era esperable en una prueba de referencia curricular.

En cuarto lugar, se ha realizado un análisis de la Validez Criterial. Como criterio de contraste se ha construido una batería de pruebas que incluyen una serie de variables consideradas predictivas de rendimiento en matemáticas (ver Cuadro 2). Estas pruebas, junto al KM-R, se administraron a una muestra de sujetos representativa y correspondientes a tres niveles académicos. A partir de los resultados de la batería de pruebas, junto con los de escala KeyMath-R y mediante el análisis Cluster, se analizan los patrones diferenciados de ejecución para cada uno de los tres niveles de los sujetos de la muestra determinándose las características o perfiles de los grupos de

Cuadro 2

ÁMBITOS EN LOS QUE SE AGRUPAN LAS VARIABLES PREDICTIVAS DE RENDIMIENTO EN MATEMÁTICAS

- Aptitudes Mentales.
- Hábitos y Estilos Cognitivos.
- Factores de personalidad.
- Factores pedagógicos.
- Rendimiento anterior.
- Ambiente socio-económico.

estudiantes. Se ha comprobado que los perfiles de los grupos resultantes se deben al nivel de rendimiento a través de todas las Subpruebas del KM-R.

Asimismo, se ha realizado un análisis de varianza de una vía (Anova) con las variables criterio entre los grupos identificados, para determinar la significación de las diferencias entre ellos.

Posteriormente, a modo de un análisis clásico de Validez Convergente, se realiza un estudio correlacional de las variables del KM-R con las variables integradas en la batería de pruebas, para determinar la significación cualitativa de las relaciones observadas y mejorar el uso diagnóstico de la prueba.

Las Anovas se distribuyen entre los grupos según los patrones de rendimiento aunque solamente algunas de las variables relativas a la evolución académica e inteligencia guardan una relación significativa para todos los grupos por el rendimiento, pudiéndose señalar las siguientes cuestiones. El volumen inferior de correlaciones significativas corresponde a los grupos con un perfil de rendimiento en el KM-R más bajo. De entre las variables criterio correspondientes a la evolución académica que muestran correlaciones significativas con los grupos por rendimiento hay que señalar las correspondientes a las variables notas obtenidas en Matemáticas y en Lengua. También hay que destacar que, los grupos de más alto rendimiento dentro del grupo de sujetos de mayor nivel académico muestran correlaciones significativas a un nivel del 99% con algunas variables de la batería de Inteligencia utilizada. Estas últimas informaciones está siendo estudiada a un mayor nivel por cuanto permitirían considerar al KM-R como una prueba de rendimiento general con un alto valor predictivo.

Insistimos en el bajo volumen de correlaciones significativas encontrado entre las variables de la batería predictiva de rendimiento y los grupos por rendimiento en el KM-R, con excepción hecha de las correlaciones mencionadas anteriormente, y en el comportamiento prácticamente aleatorio de la distribución de las restantes correlaciones significativas. Por otra parte, se han elegido como variables criterio aquellas que la investigación ha ido considerando como predictivas de rendimiento en matemáticas.

Por todo ello, podemos afirmar que los resultados encontrados en este trabajo no corroboran las conclusiones de dichas investigaciones y que el perfil del estudiante de matemáticas encontrado es diferente al esperado.

Finalmente, y con el fin de cubrir todas las necesidades que comporta todo estudio de adaptación de una prueba de evaluación del rendimiento, se ha realizado un estudio diferencial del rendimiento como base a una posible baremación y, por último, ha realizado dicha baremación.

Como conclusión del trabajo de adaptación realizado, se puede afirmar que la Escala *KeyMath-Revised* de A. Connolly puede ser considerada una prueba general de matemáticas elementales referida al currículum y con una organización interna correspondiente al currículum. Ha mostrado unos niveles de bondad métrica muy aceptables, mejorando los estudios originales de Connolly, que avalan su uso; asimismo, teniendo en cuenta los resultados obtenidos en los diferentes estudios de validación, se revela como un instrumento diagnóstico de gran utilidad.

Además, la fundamentación teórica realizada a partir de algunas conclusiones del enfoque del Procesamiento de la Información y las revisiones lógicas de la Validez de Contenido realizadas, no existentes en la adaptación original de Connolly, otorgan al KM-R una mayor riqueza de interpretación diagnóstica y permiten diseñar programas de intervención basados en el propio currículum y en las conclusiones de la propia psicología cognitiva.

REFERENCIAS

- Bachor, D.G. (1989-90). *KeyMath-Revised (KMR)*. *Diagnostique*, vol. 15, nº 1-4, pp. 87-98.
- Connolly, A.J. (1985). *KeyMath Revised: teach and practice: Activities for the diagnosis and remediation of computation difficulties*. Circle Pines, Minnesota. American Guidance Service.
- Connolly, A.J. (1988). *KeyMath Revised: A Diagnostic Inventory of Essential Mathematics*. Circle Pines, Minnesota. American Guidance Service.
- Connolly, A.J. (1988a). *KeyMath Revised Assist. Automated System for Scoring and interpreting Standardized Tests*. Circle Pines, Minnesota. American Guidance Service.
- Mari, R. (1996). *Evaluación del rendimiento en matemáticas: estudio métrico diferencial de la escuela KeyMath-Revised*. Tesis Doctoral. Dep. MIDE, Universitat de València (inédito).
- National Council of Teachers of Mathematics (1980). *An agenda for action: Recommendations for school mathematics of the 1980's*. Reston, Virginia. Autor.
- National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards for School Mathematics*. (4ª Ed. 1991). Reston, Virginia. Autor.
- Pandey, T. (1994). *Authentic Mathematics Assessment*, Paper. California Department of Education. Obtenido por acceso mediante RED el 25 de mayo de 1994.
- Price, P.A. (1984). A comparative study of the California Achievement Test (Forms C and D) and the KeyMATH Diagnostic Arithmetic Test with secondary LH students. *Journal of Learning Disabilities*, vol. 17, nº 7, pp. 392-396.

- Romberg, T.A. (Ed.) (1992). *Mathematics assesment and evaluation. Imperatives for Mathematics Educators*. New York. State University of New York Press.
- Romberg, T.A., Wilson, L.D., M. Khaketla y Chavarria, S., (1992). Curriculum and test alignment. En Romberg, T.A. (Ed.), *Mathematics assesment and evaluation. Imperatives for Mathematics Educators*. New York. State University of New York Press, pp. 61-74.
- Tinney, F.A. (1975). A comparison of the KeyMath Diagnpstic Arithmetic Test used with learning disabled students. *Journal of Learning Disabilities*, vol. 8, n° 5, pp. 57-59.